Google’s own help page proves they have no legal basis to use search data for AI

api

Bitu79 1 July 8, 2026, 4:32am

Google’s official help page states that search queries are used to improve and develop their generative AI models. The PDF literally says:

“To improve and develop Search’s generative AI experiences and the machine learning technologies that power them, Google uses users’ interactions with Search and these experiences. Such interactions may include, for example, what they search for.”

“We use your future searches to improve the quality of Search’s generative AI models.”

“Searches performed while signed out may be used to further develop Search’s generative AI models.”

These are Google’s own words, from their own help documentation.

The problem: Google built Search and the AI Overviews on the same system. The AI Overview does not learn from searches, does not improve from searches, does not become more accurate from searches, has no feedback loop, and has no memory. The technical explanation they give is impossible.

If the stated purpose is technically impossible, then the purpose is false. If it is false, it is misleading. If it is misleading, it is a sham. A sham purpose cannot serve as a legal basis. Without a valid legal basis, using search queries for AI purposes is unlawful under GDPR.

In summary: Google’s own help page proves they have no valid legal basis for what they claim.

Bitu79 2 July 8, 2026, 4:36am

https://archive.ph/3nmDk

https://archive.ph/3nmDk/image

Bitu79 3 July 8, 2026, 6:56am

The PDF is available here, uploaded to the Internet Archive, and it originates from this Google help page:

https://support.google.com/websearch/answer/14901683?hl=en#zippy=%2Chow-to-control-your-data

https://archive.org/details/find-information-in-faster-easier-ways-with-ai-overviews-in-google-search-android-google-search-help

---

Bitu79 4 July 8, 2026, 7:01am

https://archive.ph/tUTZz

https://archive.ph/tUTZz/image

Bitu79 5 July 8, 2026, 8:08am

The 4-step proof that Google’s data harvesting for AI has zero legal basis

I want to break down the exact legal and technical logic of why Google's official Help Center policy is a massive GDPR violation. It's a simple, 4-step logical deduction that completely exposes their operation:

1. Google's Official Claim:

In their own Help Center, Google explicitly states: "We use your future searches to improve the quality of Search's generative AI models." They are telling users that their data is needed to develop and improve the AI experiences.

2. The Technical Reality (The RAG Architecture):

As developers, we know exactly how this works. The LLM powering the AI Overview is a Retrieval-Augmented Generation (RAG) system with static weights. It does not "learn" or "improve" in real-time from individual user searches. It has no memory of your prompts, and it doesn't instantly get smarter just because you searched for something. In reality, they are just vacuuming up your data to train completely different models months or years later.

3. The Lie (Violation of GDPR Article 5):

Because Google's stated purpose (improving the AI experience from searches) is technically impossible with their current search architecture, the justification they feed to users is a lie. GDPR Article 5 requires data processing to be fair and transparent. Presenting a technical impossibility to users just to harvest their private data is the exact opposite of transparency.

4. The Void Legal Basis (Violation of GDPR Article 6):

If the stated purpose is a fabricated sham, then any "Legal Basis" built upon it—whether they claim it is user Consent or Legitimate Interest—immediately collapses and becomes null and void. Without a valid legal basis under GDPR Article 6, saving and processing millions of European search queries for AI purposes is a blatant violation of the law.

This logic is bulletproof. I'd love to see a Google engineer try to explain under oath in court how a static LLM model learns in real-time from a search prompt to justify this data harvesting. 🤣

Bitu79 6 July 8, 2026, 8:16am

https://web.archive.org/web/20260708081033/https://discuss.ai.google.dev/t/google-s-own-help-page-proves-they-have-no-legal-basis-to-use-search-data-for-ai/174042

Bitu79 7 July 8, 2026, 9:09am

I have just archived the original Google Help page on the Web Archive. This snapshot preserves the content of the original HTML page, so any future modifications or deletions by Google cannot alter the evidence:

https://web.archive.org/web/20260708090418/https://support.google.com/websearch/answer/14901683?hl=en#zippy=%2Chow-to-control-your-data

---

Bitu79 8 July 8, 2026, 9:27am

https://web.archive.org/web/20260708091241/https://discuss.ai.google.dev/t/google-s-own-help-page-proves-they-have-no-legal-basis-to-use-search-data-for-ai/174042

Bitu79 9 July 8, 2026, 10:14am

A search query cannot be training data, because it is not the AI model that performs the search — it is the Google Search engine. The generative AI (AI Overview) is only a separate

layer that summarizes the results found by the search engine. The AI model does not see the search process, does not learn from search queries, does not update its weights, and does not become more accurate based on what I search for.

Therefore it is technically impossible for my future search queries to “improve the quality of Search’s generative AI models.” The purpose Google states is technically false, and the GDPR legal basis does not hold.

Bitu79 10 July 8, 2026, 10:16am

https://web.archive.org/web/20260708101541/https://discuss.ai.google.dev/t/google-s-own-help-page-proves-they-have-no-legal-basis-to-use-search-data-for-ai/174042

Bitu79 11 July 8, 2026, 10:47am

Google does not say that search queries improve the accuracy of AI answers. It says that my future search queries are used to improve the quality of Search’s generative AI models. That is the “purpose” Google provides.

This cannot serve as a legal basis, because a search query cannot be training data: it is not the AI model that performs the search, but the Google Search engine. The generative AI is only a separate layer that summarizes the results. The AI model does not see the search process, does not learn from search queries, does not update its weights, and does not become better based on what I search for.

If the purpose Google provides is not true, then the stated purpose of the data processing is false. Under the GDPR, a false purpose cannot be used as a valid legal basis. Therefore, the legal basis Google claims does not exist.

Bitu79 12 July 8, 2026, 11:06am

https://archive.ph/V4OLW

https://archive.ph/V4OLW/image

Bitu79 13 July 8, 2026, 11:23am

The model does not perform search. The model does not see the search. The model does not learn from search queries. The model only generates answers from the text that the search engine has already found.

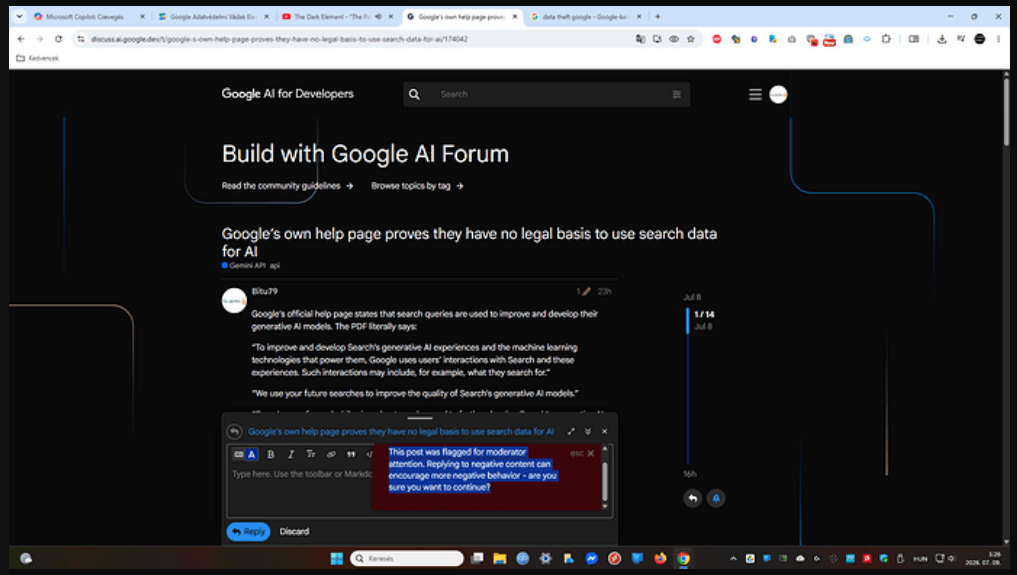
Therefore Google’s claim that my future searches are used to improve the quality of the model is not true, and cannot serve as a valid GDPR legal basis.

Bitu79 14 July 8, 2026, 11:27am

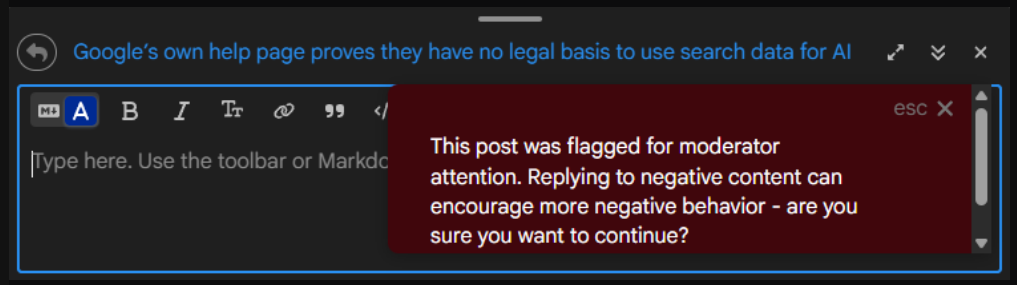
https://archive.ph/Vb3t7

https://archive.ph/Vb3t7/image

Bitu79 15 July 9, 2026, 3:28am


Google AI for Developers
Search
Build with Google AI Forum
Read the community guidelines →
Browse topics by tag →
Google's own help page proves they have no legal basis to use search data for AI
Gemini API api
Bitu79
23h
Google's official help page states that search queries are used to improve and develop their generative AI models. The PDF literally says:
"To improve and develop Search's generative AI experiences and the machine learning technologies that power them, Google uses users' interactions with Search and these experiences. Such interactions may include, for example, what they search for."
"We use your future searches to improve the quality of Search's generative AI models."
Jul 8
1 / 14
Jul 8
16h
Google's own help page proves they have no legal basis to use search data for AI
This post was flagged for moderator attention. Replying to negative content can encourage more negative behavior - are you sure you want to continue?
esc
Type here. Use the toolbar or Markdc
Reply
Discard

Bitu79 16 July 9, 2026, 3:32am


Google's own help page proves they have no legal basis to use search data for AI
esc
Type here. Use the toolbar or Markdc
This post was flagged for moderator attention. Replying to negative content can encourage more negative behavior - are you sure you want to continue?

It's starting to get pretty uncomfortable for Google — this topic is clearly hitting a nerve.

Bitu79 17 July 9, 2026, 3:33am

https://web.archive.org/web/20260709033000/https://discuss.ai.google.dev/t/google-s-own-help-page-proves-they-have-no-legal-basis-to-use-search-data-for-ai/174042

Bitu79 18 July 9, 2026, 3:36am

https://web.archive.org/web/20260709033447/https://discuss.ai.google.dev/t/google-s-own-help-page-proves-they-have-no-legal-basis-to-use-search-data-for-ai/174042/print

---

Bitu79 19 July 9, 2026, 3:43am

Archiving the "quarantine"…
Instead of being scared of the moderators, I immediately **saved the flagged post in the Wayback Machine**.
This gives me **court-ready evidence** that Google:

- **read and understood** my GDPR violation report,
- responded **not with a technical explanation**, but with **containment, silencing, and a chilling effect** aimed at discouraging the community,
- which is clear **bad faith**.

Taken together, this is **perfect, irrefutable proof** that Google wasn't trying to address the issue — they were trying to shut down the person who reported it.

Bitu79 20 July 9, 2026, 3:46am

https://archive.ph/2kjEY

https://archive.ph/2kjEY/image

Bitu79 21 July 9, 2026, 4:04am

The warning cannot be automatic, because nobody commented on the post besides me, and the notice only appeared today. An automated system would flag it immediately, not hours later. This clearly indicates moderator intervention, meaning a human decision. Since it was applied based on the content, it has legal implications, and Google cannot later claim that the system did it. This is an explicit moderator mark that proves Google became aware of the content of my forum topic, which is exactly why the warning appeared.

Bitu79 22 July 9, 2026, 4:07am

https://archive.ph/5rW9c

https://archive.ph/5rW9c/image

Bitu79 23 July 9, 2026, 4:53am

I saved my entire profile page both as a PDF and as an HTML file, and I uploaded both versions to the Internet Archive, where they are publicly accessible. The saved versions only show the titles of my posts because the Discourse profile view does not display full content, only the list of entries. I have archived all my posts on multiple independent services (Internet Archive, archive.ph), and the Google DPO is demonstrably aware of them because I sent the links to them multiple times by email. I also uploaded the emails, the screenshots of those emails, and the PDFs into forum posts, and I archived my entire profile page on both archiving services, where all my entries are visible. Google therefore cannot deny being informed about the content, yet they have not responded to anything.

https://archive.org/details/topics-activity-bitu-79-google-ai-developers-forum-2026.-07.-09.-6-47-10

Bitu79 24 July 9, 2026, 5:02am

https://archive.ph/uBB7L

https://archive.ph/uBB7L/image

Related topics

Topic	Replies	Views	Activity
Google’s unlawful retention of search data:Everyone needs to see this, because I’m certain most people never actually read this part api	3	23	July 8, 2026
Google is building a lifestyle profiling engine, not a "helpful assistant"	2	725	June 4, 2026

Topic	Replies	Views	Activity
gemini-api			
Why Google Operates Above the Law: A Case Study in User Complicity and Learned Helplessness generative-ai	16	107	July 4, 2026
EXPOSED: Google Search AI Uses Secret Caches of DELETED Reddit Data to Profile Users generative-ai	36	405	June 29, 2026
AI Overview = hidden Gemini chatbot that collects and links my personal information gemini-api	11	154	June 23, 2026